8-2017

# Using Naturalistic Data for Bicycle Safety Analysis: An Application of CyclePhilly Data to Assess Wrong-Way Riding

Nirbesh Dhakal
*University of Tennessee, Knoxville*, ndhakal@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Nirbesh Dhakal entitled "Using Naturalistic Data for Bicycle Safety Analysis: An Application of CyclePhilly Data to Assess Wrong-Way Riding." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Civil Engineering.

Christopher R. Cherry, Major Professor

We have read this thesis and recommend its acceptance:

Hyun Kim, Lee D. Han

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

# Using Naturalistic Data for Bicycle Safety Analysis: An Application of CyclePhilly Data to Assess Wrong-Way Riding

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Nirbesh Dhakal
August 2017

# ACKNOWLEDGEMENTS

I would like to thank my academic advisor, Dr. Christopher R. Cherry for his continuous guidance, support and mentorship in all the work I have done including the writing of this thesis. I would also like to thank Dr. Lee D. Han and Dr. Hyun Kim for their involvement as a part of my thesis committee and for the continuous support.

I would also like to acknowledge Southeastern Transportation Center for funding this project and Delaware Valley Regional Planning Commission for making the data available for this research. The results presented in this report do not necessarily reflect the views or position of DVRPC or STC.

Finally, I would like to thank my parents for their continuous love and support. Also, a big thank you to my sister for always motivating and being there for me. I also want to thank my fellow lab mates at Transportation Engineering and Science program, who have supported me in all occasions. This work would not have been possible without the support of all of you.

## ABSTRACT

This study focuses on wrong-way riding and associated route-choice determinants of wrong-way riding. I present this application as a case study of the many types of analysis that are possible with emerging naturalistic datasets that are not limited to conventional methods (e.g., vehicle counters). I used a dataset generated from a smartphone application, CyclePhilly, that measures rider location second-by-second in Philadelphia. The dataset covers 12,202 trips by 300 unique CyclePhilly users collected from May 2014 through April 2016. The data also includes information on socio-economics of the rider, as well as cycling experience. I merged this dataset with complementary network datasets (like speed limits and traffic levels). The data allows us to identify route choice information that includes origins, destinations, and street segments chosen. Comparing the routes travelled with the network information, allowed to assess the proportions of those trips (segments) that include wrong-way riding for each segment. From these analyses, I compared the routes with and without wrong-way riding segments and to compare the network and demographic factors that could influence wrong-way riding behavior. It was found that trips made for commute purposes were more likely to have wrong-way riding than trips made for other purposes. Different bike infrastructure showed different effect on the wrong-way riding behavior. While bike lanes and cycle tracks showed higher wrong-way riding, sharrow lanes and buffered bike lanes were found to discourage that behavior. Roads with higher average AADT also showed to have less wrong-way riding than roads with less AADT. Roads with higher number of lanes also showed more wrong-way riding. The results from this study will help engineers and planners justify the suitability of various engineering or education measures to address wrong-way riding when planning for bike facilities.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER ONE
# INTRODUCTION

Bicycling has been a popular mode of transportation used for commuting, recreation and exercise purposes. Netherlands and Denmark have a significant proportion of trips made by bicycle (30% and 20 % of all trips respectively) (Pucher, Evans et al. 1998). However, cars are the predominant mode of transport of all modes in the US. Bicycling accounts for only 1% of all trips and cars are used for 84% (Pucher, Evans et al. 1998). But bicycling is becoming increasingly popular. The number of people who commute to work by bicycle increased about 60 percent over the last decade (United States Department of Commerce. Bureau of the 2014). The trend has increased towards using active modes of transportation. Active transportation is any form of human-powered methods of travel, such as walking, cycling, using a wheelchair, or skateboarding. Active modes of transportation are environmentally friendly, have health benefits, make people social, and reduce road congestion.

While there are many benefits of bicycling, its increasing use is a concern because the inadequate biking infrastructure undermines the safety of the bicyclist. In 2014, 726 people lost their lives in bicycle/motor vehicle crashes and the number of estimated bicyclist injuries climbed to 50,000. (NHTSA 2016). Furthermore, data on bicycle crashes and safety are scarce. The major sources of these data are police reports, and many of the incidents involving bicycles go unreported (Vanparijs, Panis et al. 2015). Bicycles usage data are also hard to collect. Planning for bicycle transportation facilities requires data that measures the bicycle flows. Transportation planners are interested in answering mostly three questions: how many people are bicycling, which facilities are frequently used that could be a focus for prioritizing new investments, and if new investments have led to increased bicycling.

Though the National Household Travel Survey collects information on travel behavior of people on various modes, this survey contains more information on automobile use than for bicycle use and bicycling tends to be underreported in travel surveys (Sharp and Murakami 2005). This gap in data for bicycle usage is partly being filled by the application data collected from smartphones. The increasing use of smartphones and the decreasing cost of installing GPS have provided new sources of data that can be used to explain the travel behavior. Route tracking and fitness tracking applications are increasing in use, collecting data that can be used by planners across various cities in the US. CycleTracks in San Francisco, Cycle Atlanta in Atlanta, CyclePhilly in Philadelphia and IBIKEKNOX are some examples of applications used to track bicycle usage. Public agencies have started to use this data to help in planning for better cities. For instance, during the Fall of 2013, Oregon Department of Transportation (ODOT) became the first public agency to purchase STRAVA dataset for use in planning research purposes. Unlike traditional survey data, these data provide detailed route data for planners, which they can use to perform micro level analysis, helping them answer questions about cycling behavior. Route data also demonstrate risky riding behavior on bicyclist's part, such as wrong-way riding.

Emerging big datasets can give unparalleled insight into cyclist behaviors and exposure and ultimately help understand how behavior and infrastructure design can be simultaneously assessed to improve cycling networks and target behavioral interventions. Several factors impact bicyclist route choice and resultant exposure to traffic risks, including built environment and bike facilities (Chen and Shen 2016, Khatri, Cherry et al. 2016). Certain cyclist behaviors are apparent and can result in safety risk. Wrong direction riding on roadways (particularly one-way) can increase risk to cyclists and others. In California, one study looked at all bicycle crashes reported through California Statewide Integrated Traffic Records System (SWITRS) for year 2012 (Stimpson, Zhu et al. 2016). It reported that 11.9% of bicyclist were travelling in wrong direction prior to collision, which made up the largest share of movement preceding collision after "proceeding straight" and greater than making left turns (4.7%) and entering traffic (7.5%). Furthermore, the data showed that among accidents where bicyclists were at fault, riding on the wrong side of the road was the number one cause of accidents in Long beach area.

One of the challenges with understanding the magnitude of wrong-way riding safety behavior is the lack of exposure data. New cycling data sources are transforming the way we analyze cyclists' behavior. Until recently, little data was available on bicycling counts, and very little information was available on revealed route choice behavior or other safety-related behavior. The increasing use of GPS on smartphones and fitness tracking applications have provide an easier way to gather high resolution data on trip and route choice behavior. While these data may not be fully representative of the population, it still provides a means to explain the captured behavior. This study is one application of many that can be exploited through app-based data sources.

In this study, I used a naturalistic dataset to look at trip attributes and features of a road segment that affect the wrong direction riding of CyclePhilly users, in Philadelphia, PA. The main objective of this study is to identify the influence of trip and road attributes on wrong-way riding behavior of cyclist. Using a dataset gathered from CyclePhilly application, this study identifies wrong-way riding behavior by comparing the observed route with roadway network information from Open Street Map. In addition to the data obtained from the website, readily available online data for traffic (AADT) and speed limits are also merged with the dataset. Route choice models that have studied the impact of built environments have relied on generating a fixed number of alternatives. This study, however, focuses on the observed wrong-way riding of cyclists to determine their influence. First, this study looks at wrong-way riding at the trip level, identifying the attributes of trips or users more likely to influence wrong-way riding. Then, I focus on each segment, aggregating all wrong-way trips over the segment and looking at its relationship with road features. This study focuses on unambiguous wrong-way riding on one-way streets, but does not consider two-way streets or streets with legal two-way riding (e.g., contraflow lanes or two way cycle tracks) because of the limited spatial resolution of the GPS. The results from this study will help engineers and planners justify the suitability of various engineering or education measures to address wrong-way riding.

The following section reviews relevant literature explaining the behavior of cyclists and their associated safety risks. This is followed by a chapter explaining the data and methods used to generate results to answer our questions. I then present our results and discuss the findings and limitations of the study, followed by a concluding chapter.

# CHAPTER TWO
# LITERATURE REVIEW

## 2.1 Infrastructure Provisions

Bicycle facilities are found to have an influence on cyclist and the trips they take. Pucher and Buehler (2008) analyzed and studied national aggregate data and case studies Netherlands, Denmark and Germany, and found that providing separate cycling facilities along heavily travelled roads and intersections in combination with traffic calming on most residential neighborhoods attributed to high levels of cycling. Similarly, Moudon, Lee et al. (2005) showed a correlation between cycling and built environment, finding that trails and bike lanes increases likelihood of cycling. Among riders, commuters especially show different characteristics than other cyclists. Commuters prefer shorter routes and are less influenced by bicycle facilities like bike lanes (Dill and Gliebe 2008, Broach, Gliebe et al. 2009). They are however more sensitive to grades along their path, showing a tendency to avoid hills (Dill and Gliebe 2008, Broach, Gliebe et al. 2009). Some studies have found that cyclists are willing to travel more distance to use a bike path (Howard and Burns 2001, Tilahun, Levinson et al. 2007), Hood, Sall et al. (2013) found that infrequent cyclists valued bike lanes more than frequent cyclists.

A San Francisco report (Gajda, Sallaberry et al. 2004) reviewed before and after videotapes of 2,400 cyclists and 2,400 motorists to compare the influence of two shared lane marking designs (a “bike in a house” marking and a bike and chevron marking). They found that the bike and chevron marking encouraged cyclists to ride 8 inches farther away from the door zone; encouraged motorists to give 2 feet 3 inches more space when they were passing cyclists; and reduced the incidence of sidewalk riding by 35%.

## 2.2 Bicycle infrastructure and wrong-way riding

A few papers have studied the path taken by cyclists and their behaviors like wrong-way riding. Wrong-way riding is one of the major concerns for the safety of the cyclists and often easily overlooked (Wachtel and Lewiston 1994). Wetchal found that wrong-way riding is dangerous on all types of facilities, especially on a sidewalk. Other studies have shown mixed results on the wrong-way riding behavior. A study on shared bicycle use in Lyon, France showed that most cyclists use sidewalks, drive the wrong way up one-way streets, or use the bus/tramway lanes (Jensen, Rouquier et al. 2010). However, Hood, Sall et al. (2013) found that while cyclists prefer fewer turns and shorter distance, they will not travel the wrong way unless it saves more than four times the distance.

By comparing the behavior of cyclists before and after installing bike lanes in a neighborhood in New Orleans, Parker, Rice et al. (2013) found that the shared-lane markings and bike lanes reduce the number of wrong-way riders. Their results are complementary to the earlier findings in a 1999 Florida Study (Hunter, Stewart et al. 1999) and a 2004 San Francisco Study(Gajda, Sallaberry et al. 2004). Hunter, Stewart et al. (1999) compared bicycles lanes (BL) and wide curb lanes (WCL) based on videotapes of almost 4,600 bicyclists in the cities of Santa Barbara, CA; Gainesville, FL; and Austin, TX as the bicyclists rode through eight BL and eight WCL intersections with varying speed and traffic conditions. Overall, 5.6 percent of the bicyclists were riding the wrong-way (i.e., facing traffic). San Francisco Study (Gajda, Sallaberry et al. 2004) found that shared lane markings significantly reduced the wrong-way( by 80%) and sidewalk riding (by 35%). Another study, (Jensen, Rouquier et al. 2010) gathered data on shared bicycling system in Lyon between 2005 and 2007 when bicycle tracks were very uncommon. They reported that the cyclists pattern resembled most closely to pedestrians, suggesting behaviors of the cyclists like using sidewalks, riding the wrong way up one-way streets, and using the bus/tramway lanes as dedicated lanes.

Langford, Chen et al. (2015) compared various behaviors of conventional and electric bike riders in a e-bike sharing system in Knoxville with many hilly one way roads surrounding the system. The behavior of e-bike users was found to be similar to regular bike users: wrong-way riding was frequent for both types of bike riders. Khatri (2015) evaluated wrong-way riding behavior of bikeshare users on trips over forty different street segments in Phoenix, Arizona. Casual (short term) bikeshare users were more likely to ride against the traffic than the registered bikeshare users.

Researchers have also characterized the perceptions of cyclists riding in wrong-way directions. Rowland, Flintham et al. (2009) surveyed several bicyclists to get their experiences and reported that some riders enjoyed riding wrong-way on a one-way street where the cars are not allowed to go. Meanwhile, Daley and Rissel (2011) formed focus group in Sydney, Australia to explore images and perceptions of cycling, their potential influence on cycling and whether these views differed between regular, occasional, and non-riders. They found that wrong-way riding, among other rule breaking behaviors, generated a negative impression of cyclists and was mostly associated with bicycle couriers.

## 2.3 Wrong-way riding and Bicycle crashes

Wrong-way riding has been a concern for the safety of bicyclists and research has been conducted to quantify its risks. Schwarz, Hamann et al. (2016) examined driver behavior in response to cyclist behaviors and identified bicycling in wrong direction as one of the most common causes of crashes. One of the earlier papers published looked at bicycle-motor vehicle crashes at intersections (Wachtel and Lewiston 1994) in Palo Alto and found that bicyclists travelling wrong-way are more at risk than the right-way riding cyclists. Riding on sidewalks was another behavior that led to more crashes. The paper concluded that the wrong-way sidewalk travel is 4.5 times more dangerous than right-way sidewalk travel and that sidewalk bicycling promotes wrong-way travel.

Other studies have supported the same conclusions. Wrong-way riding either in the street or on a sidewalk is a frequent factor in bicycle-motor vehicle crashes (Hunter, Stutts et al. 1996). In another study, Harris, Reynolds et al. (2013) examined the impact of transportation infrastructure at intersection and non-intersection locations on bicycling injury risk finding that cyclists entering the intersection from sidewalks or riding in local streets were more likely to have wrong-way crashes.

Yan, Ma et al. (2011) analyzed motor vehicle–bicycle crashes using 4 years of reported crash data (2004–2007) in Beijing. They studied the influence of risk factors like bicyclist demographics, roadway geometric design, and road environment over irregular maneuvers, crash patterns and bicyclist injury severity. Angle collision was found to be the most common crash pattern associated with irregular maneuvers. They also found that the presence of a median discouraged the riding against the traffic and that bicyclists aged 46 to 65 were unlikely to ride against the traffic. They also reported more head-on collisions when the bicyclist was travelling against the traffic.

## 2.4 Analytical methods to assess big data

While various methods have been used to study riding behavior of cyclists, most work has been done with Stated Preference (SP) studies. SP studies have been used in the past to understand the different trade-offs bicyclists make in their route choice by giving side-by-side route options to choose from (Krizek and Johnson 2006, Sener, Eluru et al. 2009). Data collection in these methods are easier and the limited number of alternatives they generate simplify the model estimation. Since they ask for the preference of the participant over different alternatives, a hypothetical scenario can be introduced in the model to see the potential impact of any proposed action. But there are also some shortcomings of SP studies (Bradley 1988). With just options on paper, it is difficult to understand the familiarity of participants on features of those options and their actual preference on real facilities.

Revealed Preference (RP) methods had previously relied on asking participants to recall the routes they had taken (Aultman-Hall and Hall 1998). But the advances of GPS devices has now made it possible to observe the route choice of the travelers in detail.

GPS devices provide a non-intrusive and low-cost way to map the riders with details on origin, destination, time, routes taken and speed. RP studies have the advantage of using actual routes and network data with GPS. Menghini, Carrasco et al. (2010), Broach, Dill et al. (2012) and Hood, Sall et al. (2013) have studied the use of RP methods to model route choice of bicyclists. While these studies have the advantage of using actual routes and network data, there are some challenges to using GPS data for RP studies. The reliability and resolution of GPS depends on the device used, and the studies generally have limited choice sets, mostly based on shortest paths or other definitions of optimum paths. Additional challenges lie in cleaning the data, completing the road network, and developing a choice set generation. Broach, Gliebe et al. (2010) developed a method of route labelling that would improve both attribute variation and reasonableness of routes.

# CHAPTER THREE
# DATA AND METHODS

This study relies on data collected from diverse sources that are paired with route dataset from a route tracking smartphone application (based on Cycletracks open source base code) that measures the location of the rider. These datasets are used to develop models to assess wrong-way riding behavior of cyclists. This section first describes the study area and various sources of data used in the study. Then, the methodology used to clean the data, identify wrong-way riding and detour routes are discussed. The section ends with the discussion on the models used and their suitability to observe the wrong-way riding behavior.

## 3.1 Study Area

The application used to collect data for this study, CyclePhilly, was based in the city of Philadelphia, PA. Philadelphia is the largest city in Pennsylvania with a population of 1,567,442 as of US 2010 census. Most of the trips observed in our study pass thorough the center core of the city, which predominately consists of a one-way street network. Figure 1 shows the road features over the study area, where red lines show the one-way streets and green lines show two-way streets. The one-way street network is very important in this analysis as it is relied upon to unambiguously identify if the cyclists are riding in the wrong direction.

## 3.2 Description of Data

The data used in this study are extracted from multiple sources. The primary data containing the route information of trips was downloaded from the DVRPC website. Then, an Open Street Maps (OSM) dataset was used to create the underlying street network. These datasets were merged with other publicly available datasets for traffic counts, crashes, and speed limits on road segments. These sources are described as follows.

### 3.2.1 CyclePhilly Data

The dataset on trips and routes was recorded from a smartphone application, CyclePhilly. The dataset covers 300 unique CyclePhilly users collected from May 2014 through April 2016, with rider location measured each second. Users create a profile in the app and open the app each time they want to record the trip. The profile of riders contains basic demographic information. Table 1 shows all the attributes about the user and the trip taken contained in the dataset. Two shapefiles were used:

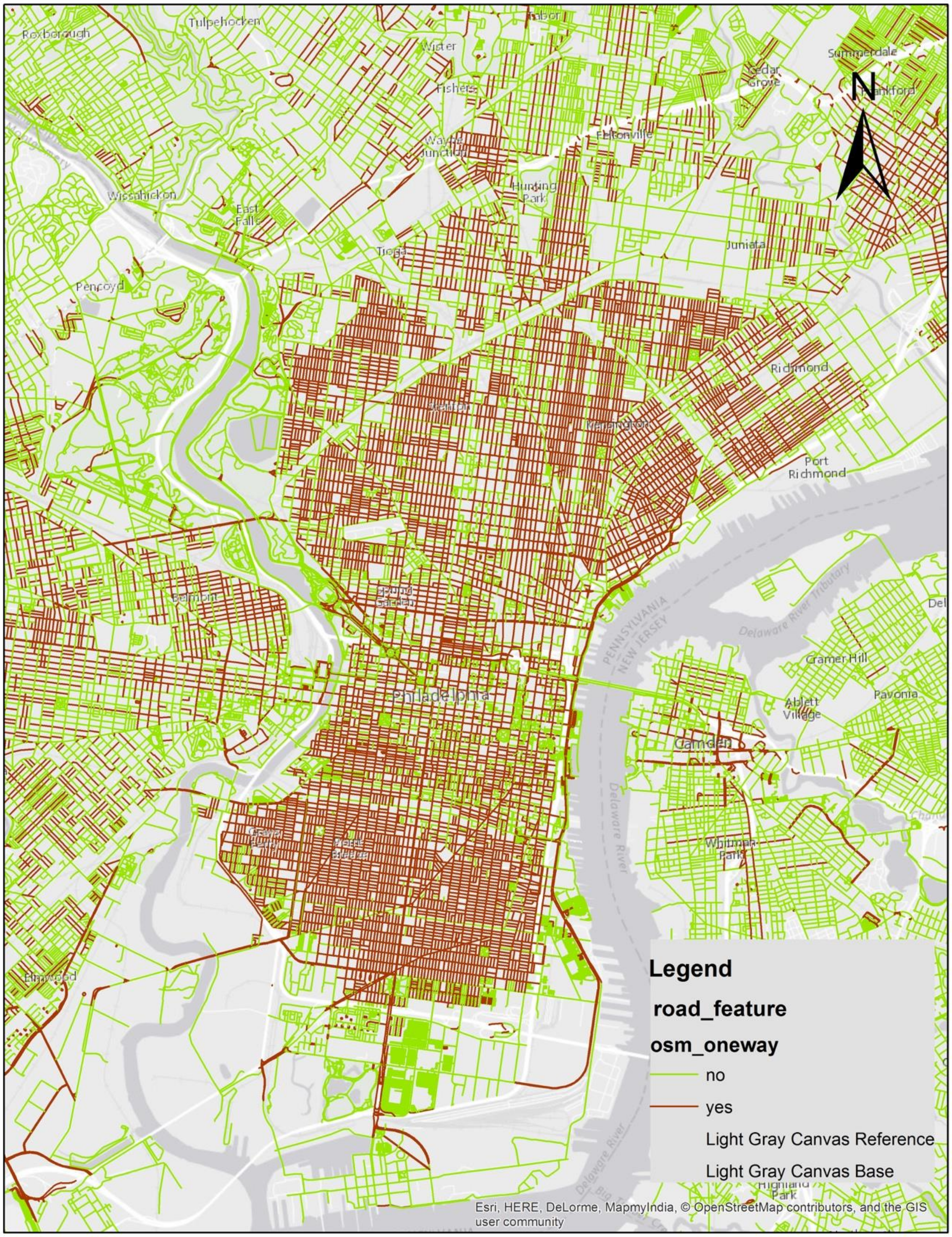


Figure 1 Figure showing one-way road segments in study area.

- Overall trip file mapped the overall trip with information of the rider.
- Segment level detail file with each trip divided into road segments.

In addition to the fields described in Table 1, the dataset also contained some information on the road network extracted from sources like OSM, DVRPC, county governments and the Bicycle Coalition. The data set contained additional information on names, length and slope of road links. DVRPC snapped the data to a 2010 version of OSM because DVRPC's transportation demand model uses that as its base network. They also made numerous in house improvements and modifications to the network.

Table 1 Description for fields in the CyclePhilly data

| Field Name | Description |
|---|---|
| UserID | User Identifier |
| TripID | Trip Identifier |
| Purpose | Trip Purpose - As selected by user |
| Start | Trip start timestamp |
| Stop | Trip end timestamp |
| Age | Age range |
| Gender | Gender |
| Income | Income range |
| Ethnic | Ethnicity |
| Cycle_freq | Cycle frequency |
| Rider_hist | Rider History |
| Rider_type | Rider Type |

### 3.2.2 OSM Street Network

The dataset downloaded from CyclePhilly was complemented with the recent Open Street Maps (OSM) data for the area. OSM data is open source data built by a community of mappers that contribute and maintain data about roads and trails. OSM data was used to build the street network for use with the network analyst extension in ArcGIS. It contains additional information on the one-way streets. For each link in the GPS data, the travelled direction was compared with this one-way direction from OSM data to check if the riders were riding in wrong direction in one way streets.

### 3.2.3 DVRPC Traffic Counts

DVRPC collects traffic volume, bicycle and pedestrian counts at over 5,000 locations each year. It also obtains traffic data collected by other entities and includes that data in its database as a public service. The data can be publicly downloaded from website opendataphilly.org. Each road segment was divided into two categories: roads with AADT higher than 3000 were classified as roads with high AADT, all other roads were classified as low AADT.

### 3.2.4 Road Speed Limit

The speed limit of each road segment was extracted from Nokia routing API. The co-ordinates for the midpoint of each road segment was used to find the speed limit category of the road. About 60% of streets have speed limit data. Streets without speed limit data were categorized as low speed roads.

## 3.3 Data cleaning

There were a couple of issues with the CyclePhilly data, mostly due to problem in the application server. A lot of the trips (around 60%) had missing values for some of the attributes for the users. Since the apps requires the user to provide their information, this might have been a reason for the missing data for the users. For the users whose information was missing, the attributes were labelled as having “no data”. Secondly, there were a lot of duplicate trips for the same users. The duplicate trips were identified by comparing the attributes of trips among each user. All the trips were checked for these two main issues to identify unique trips without missing information. The following steps were followed:

- In ArcGIS, using “Select by Attributes” feature, only the trips with values for all the attributes present were selected and exported to a different file.
- In the new file, “Dissolve” tool was used to identify two or more trips containing same time stamp for the same user. The duplicate trips identified from this process were removed from our analysis.

This process reduced the total number of valid trips to 3045 trips from 12,202 initially. The data cleaning was necessary as this study is modelling the attributes for the trips and people against wrong-way riding behavior.

## 3.4 Identifying Wrong-way Riding

After cleaning the data, the first task was to analyze the data to identify the wrong-way riding behavior in all the trips. This was done by comparing the direction of correct way riding in the one-way streets with the direction in which the rider was riding. The CyclePhilly data had the features for the trips taken up to the resolution of each segment, i.e. the direction of riding for each segment could be identified. The OSM data had the attributes for each segment, containing information if a segment was classified as a one-way street. If the segment was indeed a one-way street, the direction of the one-way street could be identified. For each segment travelled in a trip, the attributes of the direction the rider is riding and the direction of road segment were compared to check if the rider was riding in the right direction. Thus, the presence or absence of wrong-way riding for each segment in each trip was marked.

The two datasets gathered lacked any common attribute to identify a same road segment by themselves. Both of our data were polyline features and a simple “join” tool in ArcGIS couldn’t be used. Therefore, they had to be spatially joined with one another. I used the “spatial join” tool in ArcGIS to join these two features after following various processes on the dataset. Figure 2 shows the basic flowchart of the process used.

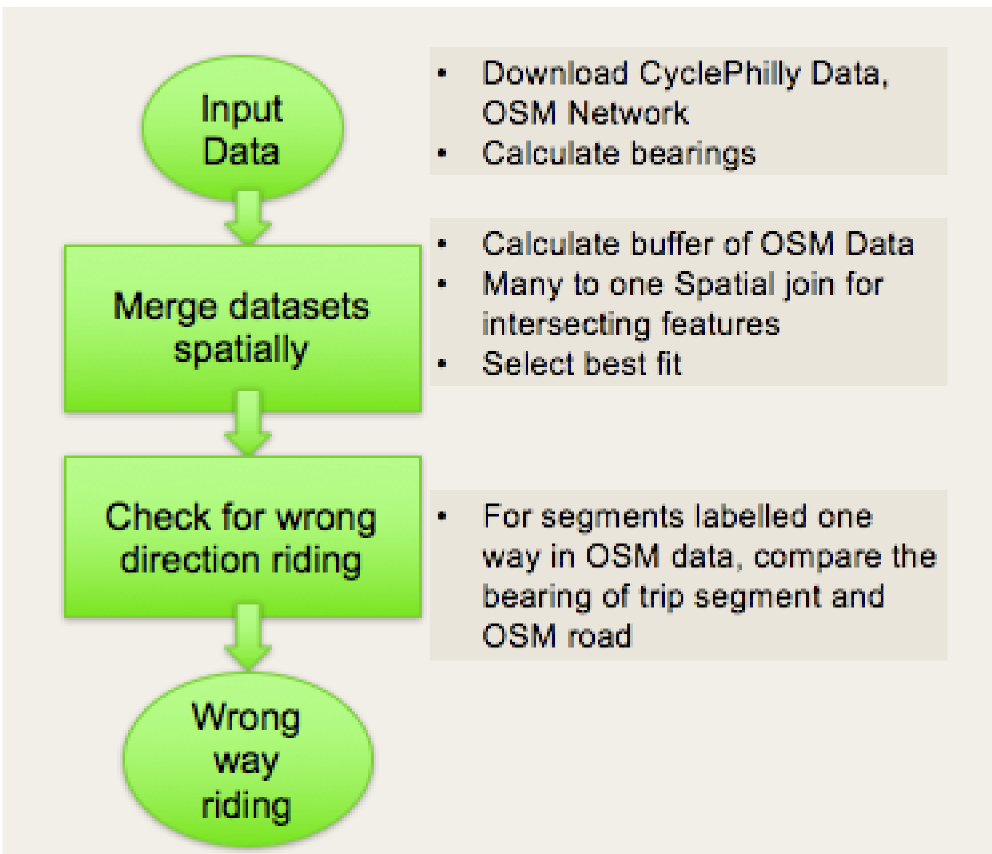


Figure 2 Flowchart showing the steps used to combine two sources of data.

The steps followed are described in detail as follows:

- First, “add bearings” from the SDM Toolbox was used to calculate the bearings for all the segments. This process was done for both datasets, and the results obtained from this process would later be used to compare those.
- A buffer of 7 meters was created around the OSM links to aid in the merge process. The distance of 7 meters was based on the road network in our study area after considering for various other distances. This step resulted in a new shape layer in ArcGIS containing all the features of the respective polyline in OSM links layer. This step made the spatial join of polyline layer of CyclePhilly data with a shape layer of OSM data easier.
- The spatial join of CyclePhilly data with the buffer layer was done in steps with various succession of “match_option” as parameter of “Spatial Join” tool. First, In the “Spatial Join” tool, CyclePhilly data (target layer) was joined to buffer layer (join layer) with “completely within” match option and “one to many” as join operation. The attributes of all the buffer layers (along with bearings for streets) which completely enclosed the CyclePhilly segment was added onto the CyclePhilly data. About 90% of the CyclePhilly trips had more than one match after this step.
- On the remaining CyclePhilly trips which were not matched to the buffer layer, “Spatial Join” tool with “intersect” match option and “one to many” join operation was used. The results of these two steps were merged in a single file which contained CyclePhilly segments associated with attributes from more than one OSM street data.
- To select the best match in between multiple matches for a single CyclePhilly segment, the bearings of the two layers joined were compared. Only the best match for each segment were retained and rest were deleted. There were some special cases, which were dealt by manually and described later in this segment.
- The resulting file contained each segment in CyclePhilly data with matched data from its respective street segment.

After the OSM attributes for each segment were found, I compared the direction of travel taken and direction of one way segments to check if the user travelled opposite to the traffic. After discovering wrong-way segments, all the segments were aggregated by Trip ID in a new file using the “Dissolve” tool in ArcGIS to determine the number of segments rode in the wrong-way in each trip. For each Trip ID, the new file contained the number of wrong direction segments and distance travelled in the wrong direction. It is important to note that the wrong-way riding was identified only on segments with unidirectional flow of traffic. For two-way streets without a median, a single line represented travel in both direction and didn’t provide enough information to identify the travel directions. Therefore those streets were not included in our analysis.

### 3.4.1 Special cases in matching two datasets

There were some special cases (around 100 observations) that had to be addressed manually. As OSM continually updates its database, making it more precise, many links are updated in the new database, adding new roads and removing old ones. Figure 3 shows some of the cases where the CyclePhilly tracks do not match the OSM data. To check how closely the two datasets were matched in the combined dataset, I compared the bearing of each line segment in both datasets. A closer examination showed that the unmatched segments were mostly found on complex intersections, like those shown in Figure 3, where the changes in road network geometry caused the difference in bearings. For those limited number of cases, I manually checked if the OSM data were correctly being matched with the Cyclephilly data.

## 3.5 Calculating lengths for detour routes

After identifying all the segments where wrong-way riding was observed, I calculated length for detour had the rider tried to avoid riding wrong-way direction. While these were not used as a variable in our models, lengths of detour routes help understand the street network. For example, if the study area is a square grid network with alternating one-way roads, we would expect a detour length to segment length ratio to be around 3:1, where the rider has to ride three blocks in the correct direction to avoid riding wrong-way on one block. For this study, detour routes were calculated for all segments where wrong-way riding behavior was observed. To calculate the detour route for those segments, a network dataset was built using OSM data. First, I downloaded a OSM file for the study area. The OSM toolbox included in ArcGIS contains tools to extract and process OSM files. After the OSM dataset was inputted using Input OSM Feature Dataset, I used the Create OSM Network Dataset tool to create the network. To assure that the routes don’t pass through roads where cyclists are not allowed to travel (eg. Freeways), I used CycleGeneric.xml network configuration file for bicycle routing.


Figure 3 Examples of difference in network in two datasets. The brown line represents the CyclePhilly data while the green line shows the OSM data.

## 3.6 Model for studying wrong-way riding behavior

Two modelling approaches were used to study the influence of various factors on wrong-way riding. First, at a trip level, I identified if the trips had wrong-way riding or not. Any trip where the rider travelled for more than 50 m in the wrong direction were classified as trips with wrong-way riding. I chose 50 m to eliminate short sub-block trip ends that might have been ambiguous, perhaps walking trips. For the binary outcome as 1 being wrong-way trips and 0 being no wrong-way trips, I looked at the attributes of the trips and the users that could influence the wrong-way riding. Next, I focus at each segment to study the influence of the attributes of the segment on the wrong-way riding behavior. Specifically, I looked at attributes of the segment like presence of bike infrastructure, AADT, speed limit, slope and number of lanes. The models used in this study for each approach is described below.

### 3.6.1 Mixed Logit model

To study the influence of user and trip attributes at a trip level, a mixed logit model was developed. In developing the logistic regression equation, the LN of the odds represents a logit transformation where the logit is the function of covariates such that

$$Y_i = \log it\ (P_i) = LN(\frac{P_i}{1-P_i}) = \ \beta_0 + \beta_1\chi_{1,i} + \beta_2\chi_{2,i} + \cdots + \ \beta_k\chi_{k,i} + \varsigma_j$$

and where $\beta_0$ is the model constant and $\beta_{0,},\ \beta_1,\ \beta_2$ are the unknown parameters corresponding with the explanatory variables (Washington, Karlaftis et al. 2010). The random intercept represents the combined effect of all omitted subject-specific covariates that causes some subjects to be more likely to ride wrong-way than others.

Figure 4 shows the general overview of the model used. The dependent variable is the model is the observed wrong-way riding behavior for a trip. The aim of our analysis was to describe in way in which wrong-way riding varies by the trip purpose, age, gender, income, cycling frequency, rider history, rider type, time of day, and length of the trip. Table 2 describes the variables used for the model that were gathered from the CyclePhilly data.

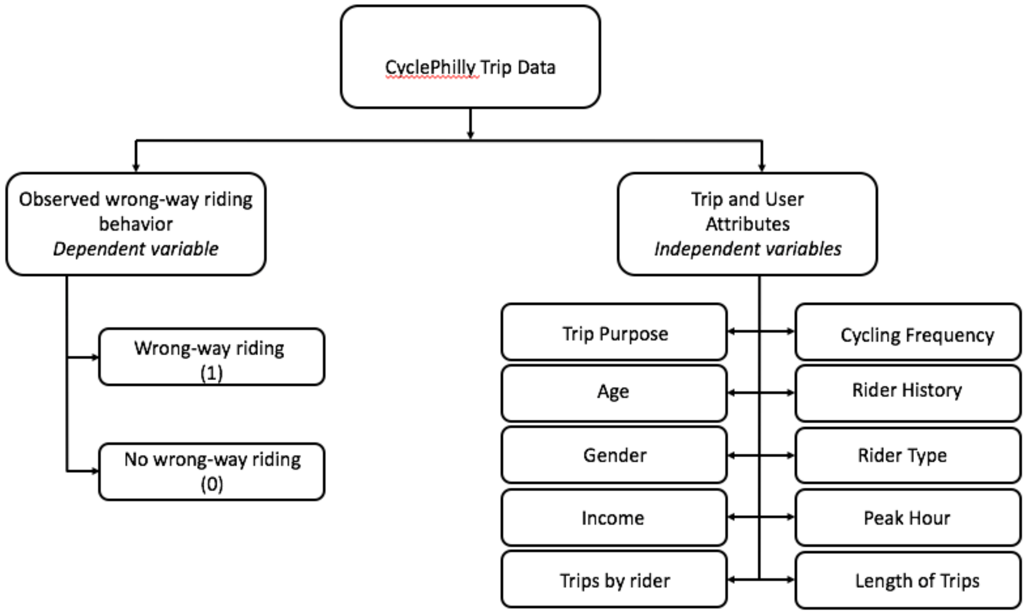


Figure 4 Mixed Logit model overview

Table 2 Independent variables in Mixed logit model

| Field Name | Description |
|---|---|
| **Purpose** | Trip Purpose - As selected by user |
| **Start** | Trip start timestamp |
| **Stop** | Trip end timestamp |
| **Age** | Age range |
| **Gender** | Gender |
| **Income** | Income range |
| **Ethnic** | Ethnicity |
| **Cycle_freq** | Cycle frequency |
| **Rider_hist** | Rider History |
| **Rider_type** | Rider Type |
| **Length** | Length of Trip |

### 3.6.2 Zero-Inflated Negative Binomial Regression Model

To estimate the influence of attributes of a segment on the wrong-way riding behavior on that segment, I used a zero-inflated negative binomial regression model. Over each one-way street segment, I accumulated the total number of trips and wrong-direction riding that were travelled over the segment. The number of wrong direction trips ridden over the segment was selected as the dependent variable in the model. About 75% of our segments had no wrong direction riding. The observation of zero events during the time period can arise from two quantitatively different conditions, failing to observe and event during the observation period, or an inability to ever experience the event.

To address this phenomenon with zero-inflected counting process, zero-inflated Poisson (ZIP) and Zero-inflected negative binomial (ZINB) regression have been developed (Washington, Karlaftis et al. 2010). ZIP model assumes that the vector of events Y=($y_1$, $y_2$,….,$y_n$) are independent and the model is:

$$y_i = 0 \; with \; probability \;\; p_i + (1 - p_i)EXP(-\lambda_i)$$

$$y_i = y \; with \; probability \;\; \frac{(1 - p_i)EXP(-\lambda_i)\lambda_i^y}{y!}$$

, where $p_i$ is the probability of being in the zero state, y is the number of events per period(number of wrong-way riding) and $\lambda$ is the expected frequency. Maximum likelihood estimates are used to estimate the parameters of ZIP regression model and confidence intervals are constructed by likelihood ratio tests.

The ZINB regression model follows a similar formulation with events, Y = ($y_1$, $y_2$, ….,
yn), being independent and model is

$$y_i = 0 \; with \; probability \;\; p_i + (1 - p_i)[\frac{1/\alpha}{(1/\alpha) + \lambda_i}]^{\frac{1}{\alpha}}$$

$$y_i = y \; with \; probability \;\; (1 - p_i) \, [\frac{\Gamma((1/\alpha) + y)\mu_i^{1/\alpha}(1 - \mu_i)^y}{\Gamma\,(1/\alpha)\, y!}]$$

where $\mu_i = (1/\alpha)/[(1/\alpha)+\lambda_i$ and $\alpha$ is regarded as over dispersion parameter, and selection of negative binomial over Poisson distribution depends on this. Γ() is a gamma function. With this model, I try to capture the effect of variables like bike specific facilities on the road, number of lanes on the road, AADT, slope, and speed limit on the road segment as described in Table 3.

Table 3 Independent variables for ZINB model

| Field Name | Description |
|---|---|
| **Bike infrastructure** | Type of bike infrastructure if present |
| **Segment length** | Length of segment |
| **Slope** | Slope of segment |
| **Crash_hist** | History of crash |
| **Speed_limit** | Posted Speed limit |
| **lanes** | No. of lanes present |
| **Total_trips** | Total trips over the segment |
| **Aadt** | DVRPC traffic counts on the segment |

Since some users took multiple trips over the same segment, I introduced a decay factor to test the sensitivity and account for the over-representation of the users on that segment. For each segment, I found the number of trips made over the segment by each user. The decay factor was selected such that the weights given to each additional trip would decrease with the increase in the number of trips. Table 4 shows the adjusted trips for the number of trips taken. The adjusted number of trips for each user riding on the segment was aggregated to find the total number of trips for each segment.

Table 4 Adjusted number of trips after decay factor

| Actual Number of Trips | Decay Factor | Adjusted Number of Trips | Adjusted Rounded |
|---|---|---|---|
| 1 | 1.000 | 1.000 | 1 |
| 2 | 1.000 | 2.000 | 2 |
| 3 | 0.667 | 2.667 | 3 |
| 4 | 0.500 | 3.167 | 3 |
| 5 | 0.400 | 3.567 | 4 |
| … | … | … | … |
| ... | ... | ... | ... |
| … | … | … | … |
| 16 | 0.125 | 5.761 | 6 |
| 17 | 0.117 | 5.879 | 6 |
| 18 | 0.111 | 5.990 | 6 |
| 19 | 0.105 | 6.095 | 6 |
| 20 | 0.100 | 6.195 | 6 |

# CHAPTER FOUR
# RESULTS AND DISCUSSION

## 4.1 Descriptive Statistics and General Results

As discussed in the data cleaning process, after removing the duplicate and erroneous trips, I was left with trips from 185 unique users. Figure 5 shows the total number of trips ridden over each segment. The trips were taken for various purposes, and commute (61%) represented the largest share of the trip purpose. Various attributes of the users were also observed. About 71% of the users were male. White repspondents represented the majority (87%) of our users in the dataset, who made 92% of all the trips. Table 5 shows the characteristics of trips for various factors observed in the study.

Table 5 Trip characteristics for various factors

| | **Variables** | **Average length (m)** | **Total trips made** | **Wrong-way trips** |
|---|---|---|---|---|
| Purpose | Commute | 4.413 | 1926 | 890 |
| | Non-commute | 3.440 | 1119 | 402 |
| Age | <34 | 4.869 | 1524 | 720 |
| | >=35 | 3.240 | 1521 | 572 |
| Gender | Male | 4.255 | 2145 | 981 |
| | Female | 3.578 | 900 | 311 |
| Income | <$20,000 | 3.731 | 192 | 113 |
| | $20,000 - $39,999 | 4.002 | 168 | 56 |
| | $40,000 - $59,999 | 3.771 | 921 | 357 |
| | $60,000 - $74,599 | 4.855 | 272 | 130 |
| | $75,000 - $99,999 | 4.695 | 636 | 340 |
| | >$100,000 | 3.715 | 856 | 296 |
| Cycling frequency | Several times a week or more | 4.320 | 2384 | 1026 |
| | Several times a month or less | 3.102 | 661 | 266 |
| Rider History | Several years or more | 4.107 | 2875 | 1209 |
| | One year or less | 3.178 | 170 | 83 |
| Rider type | Confident | 4.159 | 2453 | 1107 |
| | Cautious | 3.626 | 592 | 185 |

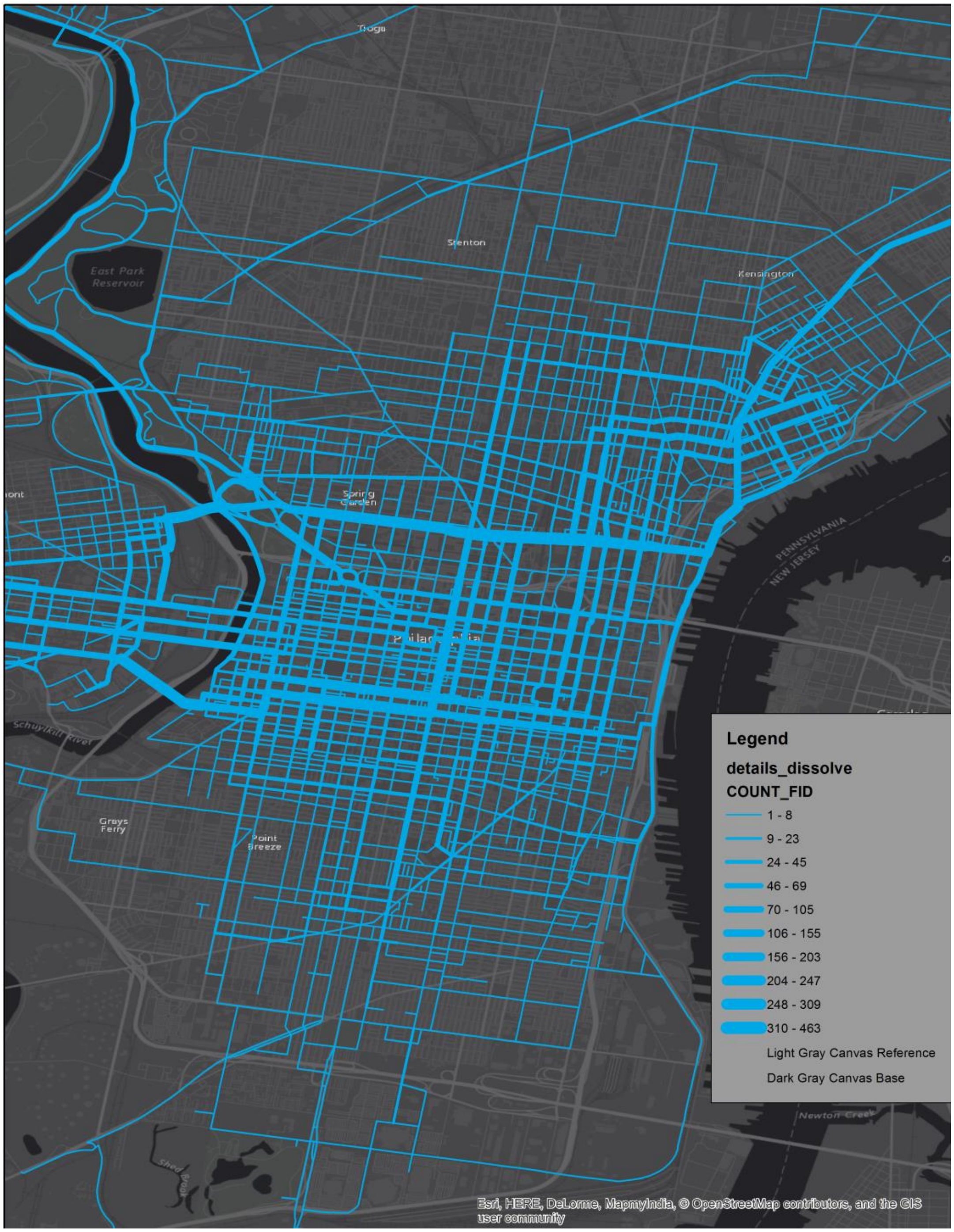


Figure 5 Total number of trips ridden over each segment

In total, the users rode 12,350 km and 2.74 % of the total distance were travelled in wrong-way direction. For the first model, I looked at the wrong-way riding over the whole trip. Among the 3,045 total trips taken, over one-third (1,292) trips had at least one wrong-way segment based on the criteria presented earlier. For the second model, I focused on each segment. The riders travelled over 4,885 unique segments. Among those segments, wrong-way riding was observed among 1,025 unique road segments. Figure 6 shows the total number of trips (in blue) and the number of trips taken in wrong-way direction (in orange) over the one-way segments, with the thickness of the line indicating the number of trips. While this is not a big number, the individual analysis of these segments helps us understand the behavior of cyclists. In addition to that, for each segment, I observed for the influence of bike infrastructures present, which is shown in Figure 7. Table 6 shows the total riding and wrong-way riding observed on each of the bike facilities present. Most of the riding was on segments with no bike infrastructure. For roads with bike facilities, segments labelled bike friendly or connectors showed higher proportion of wrong-way riding than buffered bike lanes. Figure 8 shows the proportion of bike facilities used for each trip purpose.

Table 6 Total segments travelled and wrong-way segments travelled over each one-way segment

| **Bike facility** | **Wrong-way riding** | **Total riding** |
|---|---|---|
| Sharrow Marking | 10 | 1022 |
| Bike Lane | 411 | 3652 |
| Buffered Bike Lane | 74 | 5755 |
| Trail/Sidewalk | 69 | 168 |
| Connector/Bike Friendly | 218 | 6315 |
| No Bike Lane | 1779 | 34034 |
| Total | 2561 | 50946 |


Figure 6 Total trips (blue) and wrong-way trips(orange) over each one-way segment.

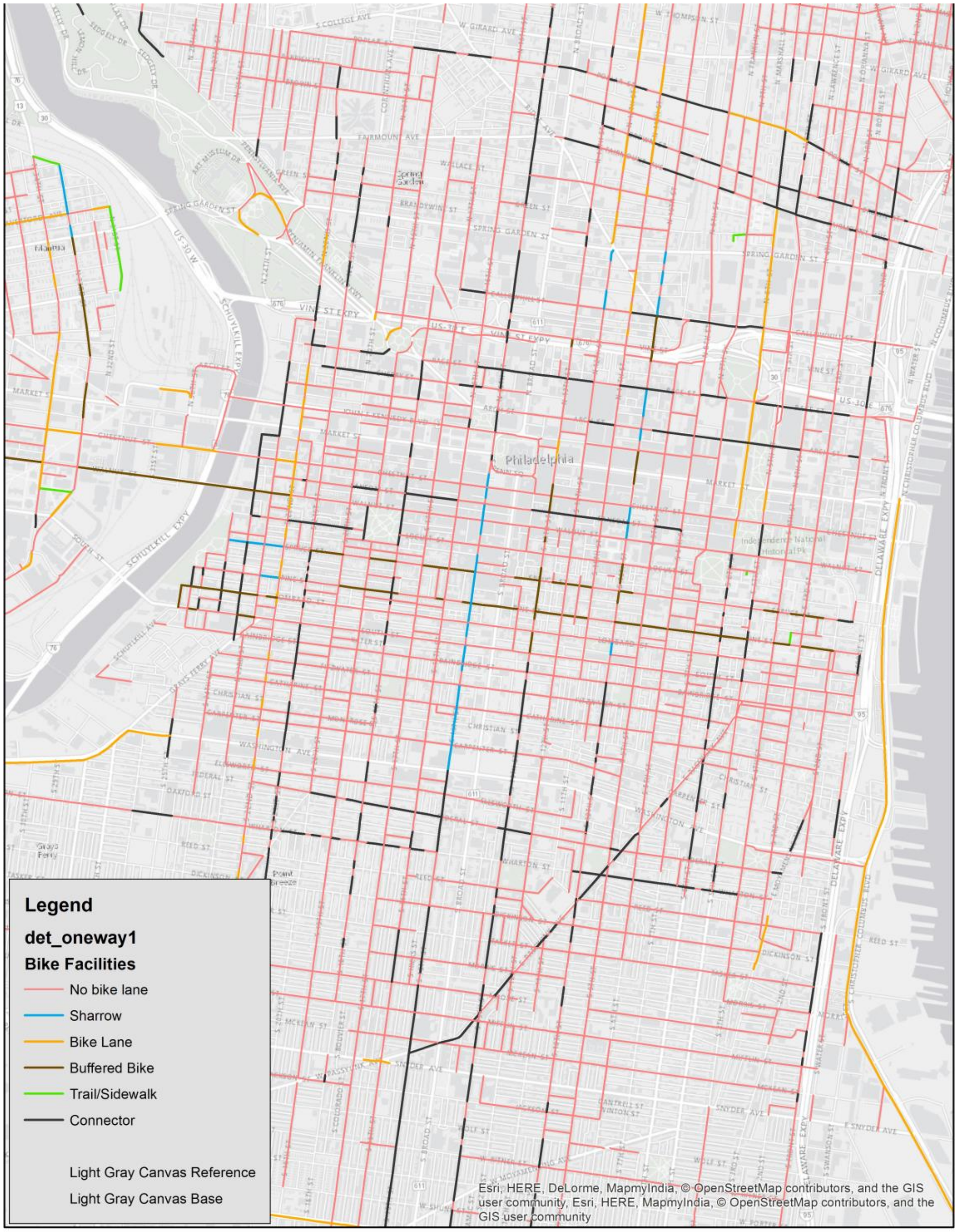


Figure 7 Bicycle facilities present over each segment.

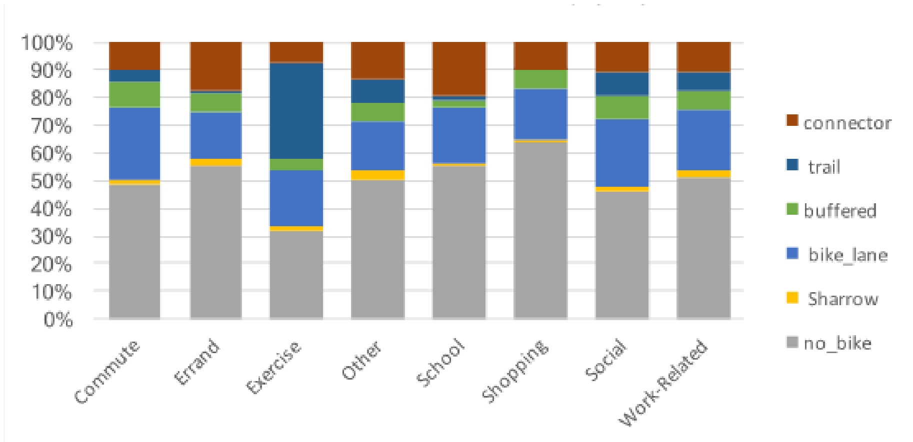


Figure 8 Proportion of bike facilities used for various purposes.

Figure 9 illustrates an example to estimate detour lengths to avoid wrong-way riding for each segment. The average ratio of detour length to the segment length was found to be 4.12. However, since the number of turns and the detour ratios were similar for detours of each segment (i.e., there was little variation), these variables were not significant in the final model and hence removed.

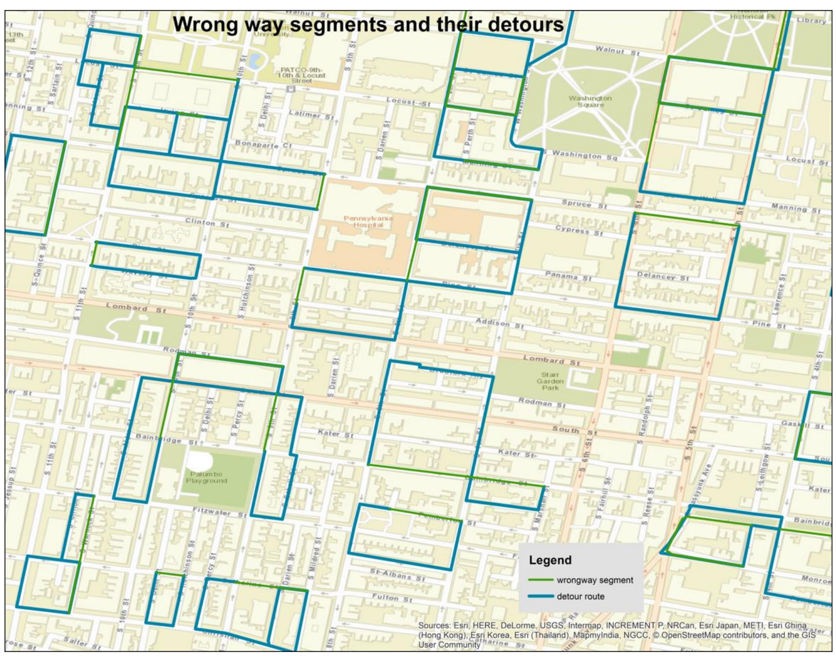


Figure 9 Figure showing detour length calculated for each segment.

### 4.1.1 Modelling for whole trips

In order to test the influence of the trip attributes on the wrong direction riding behavior, I used a mixed-effects logistic regression model. As some users made more than one trip, the data violated the IID assumption of each observation. In the mixed-effects logistic regression model, a random intercept was used which accounted for the multiple trips taken by the same user. The dependent variable was a binary function, with trips containing wrong-way riding in for more than 50m considered wrong-way trips. The results from the full model are shown below in Table 7.

Table 7 Results from Mixed effects logistic regression model

| | Full model | | | Significant only | | |
|---|---|---|---|---|---|---|
| Factors | Coef. | Std. Err. | Sig. | Coef. | Std. Err. | Sig. |
| Purpose (Commute) | 0.179 | 0.107 | 0.094 | 0.208 | 0.105 | 0.048 |
| Age | 0.075 | 0.223 | 0.737 | - | - | - |
| Ethnicity | -0.11 | 0.333 | 0.741 | - | - | - |
| Rider Type | 0.25 | 0.262 | 0.34 | - | - | - |
| Rider History | -0.42 | 0.495 | 0.396 | - | - | - |
| Peak Hour | 0.113 | 0.102 | 0.271 | - | - | - |
| Travel time | 0.004 | 0.001 | 0.023 | 0.003 | 0.001 | 0.031 |
| Trip length(km) | 0.22 | 0.021 | 0 | 0.221 | 0.021 | 0 |
| Constant | -1.287 | 0.548 | 0.019 | -1.525 | 0.154 | 0 |
| | | | | | | |
| Random-effects | Estimate | Std. Err. | | Estimate | Std. Err. | |
| UserID: Identity | | | | | | |
| sd(_cons) | 1.014 | 0.109 | | 1.0305 | 0.109 | |
| | | | | | | |
| Goodness of Fit | | | | | | |
| Log likelihood | -1725.97 | | | -1727.39 | | |
| Wald Chi-square | 131.72 | | | 128.58 | | |
| Prob. >Chi-square | 0 | | | 0 | | |
| Sample size | 3045 | | | 3045 | | |

Among all the variables tested for the model, only the purpose of the trips, trip length and travel time were significant predictors of wrong-way-riding. The other variables with attributes for the users or trips were insignificant. Compared to the trips made for commute, non-commute trips were less likely to travel in wrong direction. Having taken a trip for non-commute purposes decreases the log odds of wrong-way travel by 0.179. This result is intuitive as commute trips are mostly taken under a time constraint. In our data, the average trip speed (total distance travelled divided by trip duration) for commuters was 30% higher than non-commuters (3.59 m/s and 2.77 m/s respectively). Hence, commuters would likely gain more from saving time riding wrong direction than other users. Also, travel time and trip length were not highly correlated (correlation coefficient of 0.3). Longer trips were more likely to have wrong-way riding behavior.

For the model with only significant variables only, the purpose of the trips became significant at 95% level in the model. The influence of travel time and trip length was similar to previous model.

### 4.1.2 Modelling for wrong-way segments

In order to observe the influence of roadway characteristics on the wrong-way riding behavior of cyclists, I used a zero-inflated negative binomial model. Since most of the segments (over 70%) contained zero wrong-way riding, I used the zero-inflated negative binomial model. However, I also tried fitting a simple negative binomial model before fitting a zero-inflated negative binomial model. As a sensitivity test, I tested the models with and without the decay factor calculated as discussed in the previous section. Table 8 shows the AIC and BIC values for each model. Since the models with smaller values are better, ZINB performed better than the simple negative binomial model in both scenarios of using or not using the decay factor.

Table 8 Comparisions of diffrent models used (smaller values better)

| Model used | | AIC | BIC |
|---|---|---|---|
| Simple Negative binomial | without decay | 8307.93 | 8392.35 |
| | with decay | 7870.07 | 7954.49 |
| ZINB | without decay | 8269.74 | 8367.15 |
| | with decay | 7814.46 | 7924.85 |

Table 9 shows the results from the ZINB model with decay factor. Both scenarios (using or not using decay factor) produced similar results, so only the model with decay factor is shown. Predictors for presence of different bike infrastructure, AADT and number of lanes in the segment in the negative binomial regression model predicting wrong-way riding counts were significant. Similarly, predictors for total wrong-way counts, segment length, and AADT in the logit part predicting excessive zeros were statistically significant.

Table 9 Results from ZINB model with decay factor

| | | Coef. | Std. Err. | z | P>\|z\| |
|---|---|---|---|---|---|
| NB state | | | | | |
| | Total counts | 0.02 | 0.00 | 6.35 | 0.00 |
| | Segment length | 0.00 | 0.00 | 4.88 | 0.00 |
| | Sharrow | -1.07 | 0.46 | -2.34 | 0.02 |
| | Bike lane | 0.76 | 0.17 | 4.37 | 0.00 |
| | Buffered bike lane | -0.94 | 0.29 | -3.24 | 0.00 |
| | Trails/Sidewalk | 1.86 | 0.50 | 3.71 | 0.00 |
| | Connector | -0.29 | 0.14 | -2.17 | 0.03 |
| | No bike lanes | 0.00 | (omitted) | | |
| | AADT_high | -0.59 | 0.10 | -5.84 | 0.00 |
| | More than 1 lane | 0.51 | 0.12 | 4.15 | 0.00 |
| | constant | -0.99 | 0.10 | -10.13 | 0.00 |
| Zero State | | | | | |
| | Total counts | -0.53 | 0.13 | -4.27 | 0.00 |
| | Segment length | -0.02 | 0.01 | -3.34 | 0.00 |
| | No bike lanes | 0.09 | 0.66 | 0.14 | 0.89 |
| | AADT_high | 4.04 | 1.00 | 4.05 | 0.00 |
| | More than 1 lane | -2.41 | 2.54 | -0.95 | 0.34 |
| | constant | -0.69 | 1.32 | -0.52 | 0.60 |

Looking at results from logit part (or zero state), the log odds of observing no wrong-way riding in the segment would decrease by 0.53 for every additional trip taken in the segment. This result is intuitive as higher number of trips taken over a segment would increase the chances of observing wrong-way over the segment. Similarly, the log odds of observing wrong-way on roads on higher AADT was 4.04 times lower than on roads with lower AADT.

For the negative binomial portion of the model, different bike infrastructure was correlated with the wrong-way riding. While sharrow lane markings, buffered bike lanes, and connector roads were negatively associated with wrong-way riding, bike lanes and trails were experience more wrong-way riding than roads with no bike facilities. Segments with sharrow lanes were less likely to be travelled in wrong-way direction (expected log(count) of 1.07 lower) than roads with no bike facilities holding other variables constant. This is similar to previous findings from a study in San Francisco where presence of sharrow markings decreased wrong-way riding by 80%(Gajda, Sallaberry et al. 2004). The roads with higher AADT also seemed to discourage wrong-way riding. Roads with low AADT had 1.8 times more wrong-way riding counts than roads with higher AADT. This is also intuitive as riders might feel unsafe when riding wrong-way on a busier street, hence discouraging the behavior. Trails or cycle tracks also showed more wrong-way riding than road with no bike lanes. Cycle tracks are physically separated from motor traffic and distinct from the sidewalk, with a separated path and the on-street infrastructure like that of a conventional bike lane. Similarly, the number of lanes also showed the positive relationship with wrong direction riding. Roads with more than one segment has an expected count 1.66 times higher than segments with a single lane. A possible explanation for this may be that for wider roads, riders would prefer not to cross to ride on the correct side of the road if they're already present on the wrong side.

# CHAPTER FIVE
# CONCLUSIONS AND RECOMMENDATIONS

The main objective of this study was to demonstrate an application of using a naturalistic data for bicycle safety analysis. This study focuses on highlighting the wrong-way riding behavior of cyclists in Philadelphia using a crowdsourced data gathered from an application. This study is unique because it is the first study focusing on wrong-way riding behavior of cyclists using a naturalistic dataset and is among the first to explore city-wide aberrant behavior using probe data.

The results from the study will help planners and engineers better plan new bike infrastructure in cities. Segments with a higher number of bike trips showed more wrong-way riding. This could make the case for contra-flow bike lanes on cities like Philadelphia with many one-way streets and high bike traffic. Contra-flow bike lanes increase connectivity in the network for cyclists, and could improve safety. The results also show the influence of various bike infrastructure on the wrong-way riding behavior of cyclists which will help the engineers in choosing between various type of bike infrastructure. In addition to this, the data used in this study can be further used to study other route choice behaviors of the cyclists with the traditional route choice modeling.

However, there are some limitations of this study. The main limitation is that the CyclePhilly dataset is not a random dataset and not representative of cyclist population of the whole city. Hence any findings of this study will only reflect the nature of CyclePhilly users and their travel behavior. While this dataset is not representative of the entire cyclist population, a dataset like this provides a valuable resource in accessing the roadway infrastructure with the resolution of data it provides. In addition to that, I have complemented those data with various other publicly available dataset of traffic counts, bicycle crashes history and road speed limit. I also accounted for a single rider being overrepresented in a road segment. For each segment, I introduced a decay factor that reduces the influence of a single rider making multiple trips.

When using GPS data for bicycle routing, due to inaccuracies associated with the GPS devices, it is hard to accurately plot the paths taken over the road segment and the side of road where the trips take place. This issue limits accurately finding the wrong direction riding for bi-directional roads. Our study area in Philadelphia is full of densely connected grid network of one way streets. This gave us a unique opportunity to correctly identify the wrong direction riding.

I present this study as an application of methods that can be used to exploit these types of dataset. Furthermore, open source data from OSM provides segment level detail of road infrastructure and can complement the collected data. In this paper, I highlighted an application of using these data to study the riding behaviors of the cyclists.

# LIST OF REFERENCES

Aultman-Hall, L. and F. L. Hall (1998). "Ottawa-Carleton commuter cyclist on-and off-road incident rates." Accident Analysis & Prevention **30**(1): 29-43.
Bradley, M. (1988). "Realism and adaptation in designing hypothetical travel choice concepts." Journal of Transport Economics and Policy: 121-137.
Broach, J., J. Dill and J. Gliebe (2012). "Where do cyclists ride? A route choice model developed with revealed preference GPS data." Transportation Research Part A: Policy and Practice **46**(10): 1730-1740.
Broach, J., J. Gliebe and J. Dill (2009). "Development of a Multi-class bicyclist route choice model using revealed preference data."
Broach, J., J. Gliebe and J. Dill (2010). "Calibrated labeling method for generating bicyclist route choice sets incorporating unbiased attribute variation." Transportation Research Record: Journal of the Transportation Research Board(2197): 89-97.
Chen, P. and Q. Shen (2016). A GPS-Based Analysis of Built Environment Influences on Bicyclist Route Preferences. Transportation Research Board 95th Annual Meeting.
Daley, M. and C. Rissel (2011). "Perspectives and images of cycling as a barrier or facilitator of cycling." Transport policy **18**(1): 211-216.
Dill, J. and J. Gliebe (2008). "Understanding and measuring bicycling behavior: A focus on travel time and route choice."
Gajda, O., M. Sallaberry and P. Tannen (2004). "San Francisco's Shared Lane Pavement Markings: Improving Bicycle Safety." San Francisco Department of Parking & Traffic: 10-16.
Harris, M. A., C. C. Reynolds, M. Winters, P. A. Cripton, H. Shen, M. L. Chipman, M. D. Cusimano, S. Babul, J. R. Brubacher and S. M. Friedman (2013). "Comparing the effects of infrastructure on bicycling injury at intersections and non-intersections using a case–crossover design." Injury Prevention: injuryprev-2012-040561.
Hood, J., E. Sall and B. Charlton (2013). "A GPS-based bicycle route choice model for San Francisco, California." Transportation letters.
Howard, C. and E. Burns (2001). "Cycling to work in Phoenix: route choice, travel behavior, and commuter characteristics." Transportation Research Record: Journal of the Transportation Research Board(1773): 39-46.
Hunter, W., J. Stewart and J. Stutts (1999). "Study of bicycle lanes versus wide curb lanes." Transportation Research Record: Journal of the Transportation Research Board(1674): 70-77.
Hunter, W. W., J. R. Stewart, J. C. Stutts, H. H. Huang and W. E. Pein (1999). "A comparative analysis of bicycle lanes versus wide curb lanes: Final report."
Hunter, W. W., J. C. Stutts, W. E. Pein and C. L. Cox (1996). Pedestrian and bicycle crash types of the early 1990's.

Jensen, P., J.-B. Rouquier, N. Ovtracht and C. Robardet (2010). "Characterizing the speed and paths of shared bicycle use in Lyon." Transportation research part D: transport and environment **15**(8): 522-524.
Jensen, P., J.-B. Rouquier, N. Ovtracht and C. Robardet (2010). "I." Transportation research part D: transport and environment **15**(8): 522-524.
Khatri, R. (2015). "Modeling Route Choice of Utilitarian Bikeshare Users from GPS Data."
Khatri, R., C. R. Cherry, S. S. Nambisan and L. D. Han (2016). "Modeling route choice of utilitarian bikeshare users with GPS data." Transportation Research Record: Journal of the Transportation Research Board(2587): 141-149.
Krizek, K. J. and P. J. Johnson (2006). "Proximity to trails and retail: effects on urban cycling and walking." Journal of the American Planning Association **72**(1): 33-42.
Langford, B. C., J. Chen and C. R. Cherry (2015). "Risky riding: Naturalistic methods comparing safety behavior from conventional bicycle riders and electric bike riders." Accident Analysis & Prevention **82**: 220-226.
Menghini, G., N. Carrasco, N. Schüssler and K. W. Axhausen (2010). "Route choice of cyclists in Zurich." Transportation research part A: policy and practice **44**(9): 754-765.
Moudon, A. V., C. Lee, A. D. Cheadle, C. W. Collier, D. Johnson, T. L. Schmid and R. D. Weather (2005). "Cycling and the built environment, a US perspective." Transportation Research Part D: Transport and Environment **10**(3): 245-261.
NHTSA. (2016). "Traffic safety Facts." from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812282.
Parker, K. M., J. Rice, J. Gustat, J. Ruley, A. Spriggs and C. Johnson (2013). "Effect of bike lane infrastructure improvements on ridership in one New Orleans neighborhood." Annals of behavioral medicine **45**(1): 101-107.
Pucher, J. and R. Buehler (2008). "Making cycling irresistible: lessons from the Netherlands, Denmark and Germany." Transport Reviews **28**(4): 495-528.
Pucher, J., T. Evans and J. Wenger (1998). "Socioeconomics of urban travel: evidence from the 1995 NPTS." Transportation Quarterly **52**(3).
Rowland, D., M. Flintham, L. Oppermann, J. Marshall, A. Chamberlain, B. Koleva, S. Benford and C. Perez (2009). Ubikequitous computing: designing interactive experiences for cyclists. Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, ACM.
Schwarz, C., C. J. Hamann and P. Scholar (2016). "Examination of driver behavior in response to bicyclist behaviors."
Sener, I. N., N. Eluru and C. R. Bhat (2009). "An analysis of bicycle route choice preferences in Texas, US." Transportation **36**(5): 511-539.
Sharp, J. and E. Murakami (2005). "Travel surveys: Methodological and technology-related considerations." Journal of transportation and statistics **8**(3): 97-113.

Stimpson, J. P., H. Zhu and F. A. Wilson (2016). "Bicyclists found at fault for bicycle crashes in California." The American journal of emergency medicine.
Tilahun, N. Y., D. M. Levinson and K. J. Krizek (2007). "Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey." Transportation Research Part A: Policy and Practice **41**(4): 287-301.
United States Department of Commerce. Bureau of the, C. (2014). American Community Survey, 2008-2012 [United States]: Public Use Microdata Sample: Artist Extract, Inter-university Consortium for Political and Social Research (ICPSR) [distributor].
Vanparijs, J., L. I. Panis, R. Meeusen and B. de Geus (2015). "Exposure measurement in bicycle safety analysis: A review of the literature." Accident Analysis & Prevention **84**: 9-19.
Wachtel, A. and D. Lewiston (1994). "Risk factors for bicycle-motor vehicle collisions at intersections." ITE Journal(Institute of Transportation Engineers) **64**(9): 30-35.
Washington, S. P., M. G. Karlaftis and F. Mannering (2010). Statistical and econometric methods for transportation data analysis, CRC press.
Yan, X., M. Ma, H. Huang, M. Abdel-Aty and C. Wu (2011). "Motor vehicle–bicycle crashes in Beijing: Irregular maneuvers, crash patterns, and injury severity." Accident Analysis & Prevention **43**(5): 1751-1758.

# VITA

Nirbesh Dhakal was born in Kathmandu, Nepal. After completing his Bachelors in Civil Engineering from Institute of Engineering, Pulchowk, he came to United States to pursue his higher studies. As a graduate student at University of Tennessee, Nirbesh worked as a graduate teaching assistant in Transportation Engineering lab. He also worked as graduate research assistant in transportation planning, and sustainability projects under the guidance of Dr. Christopher Cherry. After graduation, Nirbesh plans to work in the consulting industry as a Transportation Engineer.